



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2016

---

## **Predicting stochastic community dynamics in grasslands under the assumption of competitive symmetry**

Lohier, Théophile ; Jabot, Franck ; Weigelt, Alexandra ; Schmid, Bernhard ; Deffuant, Guillaume

**Abstract:** Community dynamics is influenced by multiple ecological processes such as environmental spatiotemporal variation, competition between individuals and demographic stochasticity. Quantifying the respective influence of these various processes and making predictions on community dynamics require the use of a dynamical framework encompassing these various components. We here demonstrate how to adapt the framework of stochastic community dynamics to the peculiarities of herbaceous communities, by using a short temporal resolution adapted to the time scale of competition between herbaceous plants, and by taking into account the seasonal drops in plant aerial biomass following winter, harvesting or consumption by herbivores. We develop a hybrid inference method for this novel modelling framework that both uses numerical simulations and likelihood computations. Applying this methodology to empirical data from the Jena biodiversity experiment, we find that environmental stochasticity has a larger effect on community dynamics than demographic stochasticity, and that both effects are generally smaller than observation errors at the plot scale. We further evidence that plant intrinsic growth rates and carrying capacities are moderately predictable from plant vegetative height, specific leaf area and leaf dry matter content. We do not find any trade-off between demographical components, since species with larger intrinsic growth rates tend to also have lower demographic and environmental variances. Finally, we find that our model is able to make relatively good predictions of multi-specific community dynamics based on the assumption of competitive symmetry.

DOI: <https://doi.org/10.1016/j.jtbi.2016.03.043>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-126034>

Journal Article

Accepted Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Lohier, Théophile; Jabot, Franck; Weigelt, Alexandra; Schmid, Bernhard; Deffuant, Guillaume (2016). Predicting stochastic community dynamics in grasslands under the assumption of competitive symmetry. *Journal of Theoretical Biology*, 399:53-61.

DOI: <https://doi.org/10.1016/j.jtbi.2016.03.043>

# Predicting stochastic community dynamics in grasslands under the assumption of competitive symmetry

**Théophile Lohier, Franck Jabot, Alexandra Weigelt, Bernhard Schmid & Guillaume  
Deffuant**

**Lohier, T.** (theophile.lohier@gmail.com)<sup>1</sup>

**Jabot, F.** (Corresponding author, franck.jabot@m4x.org)<sup>1</sup>

**Weigelt, A.** (alexandra.weigelt@uni-leipzig.de)<sup>2</sup>

**Schmid, B.** (bernhard.schmid@ieu.uzh.ch)<sup>3</sup>

**Deffuant, G.** (guillaume.deffuant@irstea.fr)<sup>1</sup>

<sup>1</sup>LISC - Laboratoire d'Ingénierie pour les Systèmes complexes, IRSTEA, 9 avenue Blaise  
Pascal, CS 20085, 63178 Aubière, France.

<sup>2</sup>Institute of Biology, University of Leipzig, Johannisallee 21-23, 04103, Leipzig, Germany.

<sup>3</sup>Institute of Evolutionary Biology and Environmental Studies, University of Zurich,  
Winterthurerstrasse 190, CH-8057, Zurich, Switzerland.

Author contributions:

TL and FJ designed the model and the analyses. AW and BS provided the data. TL performed  
the analyses. TL and FJ wrote the manuscript with inputs from all authors.

Running head: biomass-based stochastic community dynamics

## ABSTRACT

Community dynamics is influenced by multiple ecological processes such as environmental spatiotemporal variation, competition between individuals and demographic stochasticity. Quantifying the respective influence of these various processes and making predictions on community dynamics require the use of a dynamical framework encompassing these various components. We here demonstrate how to adapt the framework of stochastic community dynamics to the peculiarities of herbaceous communities, by using a short temporal resolution adapted to the time scale of competition between herbaceous plants, and by taking into account the seasonal drops in plant aerial biomass following winter, harvesting or consumption by herbivores. We develop a hybrid inference method for this novel modelling framework that both uses numerical simulations and likelihood computations. Applying this methodology to empirical data from the Jena biodiversity experiment, we find that environmental stochasticity has a larger effect on community dynamics than demographic stochasticity, and that both effects are generally smaller than observation errors. We further evidence that plant intrinsic growth rates and carrying capacities are moderately predictable from plant vegetative height, specific leaf area and leaf dry matter content. We do not find any trade-off between demographical components, since species with larger intrinsic growth rates tend to also have lower demographic and environmental variances. Finally, we find that our model is able to make relatively good predictions of multi-specific community dynamics based on the assumption of competitive symmetry.

## KEYWORDS

demographic stochasticity; environmental stochasticity; ecological modelling; plant functional trait; statistical inference

## 1. Introduction

Plant community dynamics is driven by intra- and interspecific interactions, and by environmental factors such as climatic conditions or soil composition. The way these processes influence community dynamics is of utmost importance for understanding community assembly (Ackerly 2003; Ejrnaes et al. 2006; Chase 2010), productivity (Mouquet et al. 2002) and stability (Sprugel 1991; Hector et al. 2010). Studies based on static descriptors of community structure have provided tests for predictions of ecological theories (Stubbs & Wilson 2004; Cornwell et al. 2006; Norden et al. 2009; Gonzalez et al. 2010). The next step is to quantitatively relate empirical observations to the underlying dynamical processes (Jabot & Chave 2011; de Mazancourt et al. 2013). In this vein, a growing number of studies aim at building dynamical models of community dynamics based on explicit ecological processes, and at calibrating these models with field data (Lande et al. 2003; Beaumont 2010), thereby belonging to the more general trend towards a more predictive ecology (Mouquet et al. 2015).

To build models of plant community dynamics, the framework of stochastic population dynamics is particularly appealing (Lande et al. 2003). This approach consists in modelling the joint effects of competitive interactions, demographic and environmental stochasticities on community dynamics. Although the role of environmental stochasticity in community dynamics has been recognized for a long time (Chesson & Warner 1981), it has been neglected in many recent analyses by community ecologists (Chisholm et al. 2014; Kalyuzhny et al. 2015). There is therefore a renewing interest in better taking into account this component of community dynamics in dynamical models (de Mazancourt et al. 2013, Kalyuzhny et al. 2015).

This general framework has been mainly applied to easily countable organisms, such as animals (Lande et al. 2003) or trees (Chisholm et al. 2014, Kalyuzhny et al. 2015). To be applied to herbaceous plants, it has been proposed to model the dynamics of plant biomass instead of population sizes (de Mazancourt et al. 2013). But two additional specificities of herbaceous plant communities have been mainly overlooked in previous studies. First, the time scale of variation in competition between plants is short, due to the temporal variability of resources and to the rapid modification of vertical community structure following plant growth (Wilson and Tilman 1993, Silvertown et al. 2015). Second, herbaceous plant communities face frequent major disturbance events leading to sudden aerial biomass drops, such as winter mortality of aerial plant tissues or agricultural harvests by mowing or grazing (Jouven et al. 2006, Jabot and Pottier 2012). These disturbance events periodically reset aerial biomass to low levels, and therefore need to be taken into account in dynamical models of community dynamics.

The present study aims at developing a model of stochastic dynamics for herbaceous plant communities based on biomass rather than population sizes, and at taking into account both the short temporal scale of between plant competition and the frequent biomass drops encountered by herbaceous plant communities. We detail an inference method to calibrate the daily time step model parameters from biomass measurements in the field at seasonal time steps, coming from the Jena biodiversity experiment (Weigelt et al. 2010). This methodological development enables us to answer to four questions on herbaceous plant community dynamics: 1) what is the respective influence of demographic and environmental variabilities on community dynamics? 2) is there an equalizing trade-off between species intrinsic growth rates and their temporal stability as would be expected for species coexistence (Chesson 2000)? 3) are species demographical characteristics predictable from

plant functional traits? 4) are multi-specific community dynamics predictable from species  
individual dynamics?

## 2. METHODS

### 2.1. A new biomass-based model of plant stochastic community dynamics

The model describes the dynamics of species aboveground biomass within a growing season and the way competition, demographic and environmental stochasticities affect species growth. In the following, we call "season" a temporal period during which plants are growing without experiencing a strong biomass decrease or removal (due to winter, harvests or consumption by herbivores). At the end of each season, the aboveground biomass is assumed to drop and this reduced biomass is used to initialize the species dynamics in the following season (Figure 1). A community is here defined as a group of plants of the same or of different species growing in the same plot  $p$ . The intra-seasonal dynamics of plant growth is modelled with a daily time step. The biomass of species  $i$  during season  $T$  in plot  $p$  after  $t$  days of growth,  $B_i(t, p, T)$ , is modelled with the following difference equation:

$$B_i(t+1, p, T) = B_i(t, p, T) + \left( r_{mi} + \sigma_{ei} u_{ei}(T) + \frac{\sigma_{di} u_{di}(T, p)}{\sqrt{B_i(0, p, T)}} \right) \cdot B_i(t, p, T) \cdot \left( 1 - \frac{B_i(t, p, T) + \sum_{j \neq i} \alpha_{ij} B_j(t, p, T)}{K_i} \right) \quad (\text{eq.1})$$

where  $r_{mi}$  is the intrinsic growth rate of species  $i$ ,  $K_i$  is its carrying capacity and  $\alpha_{ij}$  is the inter-specific competition coefficient describing the effect of species  $j$  on species  $i$ . Environmental stochasticity encompasses the inter-seasonal variability in species growth rates stemming from environmental variability. It is modelled through  $\sigma_{ei} u_{ei}(T)$  where  $\sigma_{ei}^2$  is the environmental variance, and  $u_{ei}(T)$  are drawn from a normal distribution with zero mean and unit variance. For each species  $i$ ,  $u_{ei}(T)$  are assumed to be constant across all plots  $p$  during

season  $T$ . Demographic stochasticity encompasses intra-specific variability. It is incorporated through  $\sigma_{di}u_{di}(T,p)$  where  $\sigma_{di}^2$  is the demographic variance, and  $u_{di}(T,p)$  are drawn from a normal distribution with zero mean and unit variance, and are thus assumed to vary across the plots  $p$ . Consequently, the growth rate of a given species is likely to differ across plots because of demographic stochasticity and across seasons because of both demographic and environmental stochasticities. The last logistic term represents intra- and inter-specific competition for resources which decreases the plant growth rate. This growth reduction due to competition increases as plants grow during the season.

The biomass at the start of the growing season  $B_i(0,p,T)$  is used in the scaling of the demographic variance in equation (eq.1) as being a proxy of the number of growing individuals in season  $T$ .  $B_i(0,p,T)$  is assumed to be constant across seasons and equal to  $B_0 = 2 \text{ gm}^{-2}$  for monocultures while in species mixtures,  $B_0$  is equally divided between species. This simplifying assumption is justified by the fact that the establishment of monoculture plots was generally well-advanced when biomass monitoring started (Roscher et al. 2004), so that biomass at the start of the growing season is primarily controlled by the height of mowing that was constant during the experiment. In species mixtures,  $B_0$  is equally divided between species. This simplifying assumption implies that we neglect possible competitive exclusions taking place progressively during the course of the experiment that would cause a progressive shift in relative covers between species in the plots, and consequently in initial relative biomass at the start of the growing season. We performed preliminary analyses considering the alternative assumption that  $B_i(0,p,T)$  is proportional to  $B_i(\text{tend},p,T-1)$ . We found that models based on this assumption failed to reproduce successive events of disappearance and recovery of some species that were observed in the experiment (data not shown). Still, intermediary assumptions of partial correlations between relative biomasses at the end of a

growing season and biomasses at the start of the following season might be further studied in  
the future.

Process stochasticity (both demographic and environmental) is included in the model through its effect on biomass growth rate  $r_m$ , following the previous work of de Mazancourt et al. (2013). This model choice is the most straightforward, since variability in environmental conditions and intraspecific variability are known to affect plant growth rates (Raven et al. 2005). An alternative model choice would have been to make process noise enter in the competition term. Although environmental conditions do affect interactions between plants (e.g., Maestre et al. 2009), we contend that it is a second order effect compared to their primary influence on plant growth rates. Besides, we assume in the model that environmental and demographic variables are constant during a season. This means that they represent the average seasonal suitability for each species population. In the same vein, we assume that stochastic variables are independent across seasons. Indeed, there is no correlation between climatic variables in successive seasons, so that this is reasonable to assume that there is no correlation between the suitability of successive seasons. These model choices are economical ways to model variability across seasons, and we will show below that this relatively simple model does successively pass the model checking tests. Finally, we assume that there is no environmental variation among the different plots, so that differences between plots are solely attributable to demographic variability. This assumption is reasonable, since the among plot variability in soil characteristics is negligible, so that environmental variability is essentially due to climatic variability which is likely to be extremely similar among plots due to the limited spatial extension of the experiment.

## 2.2. Field data



We used data from the long-term grassland biodiversity experiment in Jena (Weigelt et al. 2010). In this dataset, 60 plant species were grown in monocultures from 2002 to 2010, with each monoculture replicated in 2 plots. In each monoculture plot, biomass was measured twice a year in June and September, in 2 samples per plot. Half of the monocultures (i.e 60 plots) were given up in 2006. Additionally, 16 mixtures of 2, 4 and 8 plant species, 14 mixtures of 16 plant species and 4 mixtures of all 60 plant species were grown from 2002 to 2008. In each polyculture plot, biomass of all plant species was measured twice a year in June and September, in 4 samples per plot. After each biomass measurement, both monoculture and polyculture plots were mown (Weigelt et al. 2010). We therefore consider in the following that there are two growing seasons per year, from April to June, and from July to September (Figure 1). Data can be accessed on the Pangaea database (<http://doi.pangaea.de/10.1594/PANGAEA.846321>)

### *2.3. Inference based on monoculture data*

The daily time step model contains four parameters per species when applied to monoculture data ( $r_{mi}$ ,  $K_i$ ,  $\sigma_{ei}$ ,  $\sigma_{di}$ ) as well as environmental and demographic seasonal variables  $u_{ei}(T)$  and  $u_{di}(T,p)$ . We present in the following how all these parameters and variables can be estimated using field biomass data of species grown in monoculture for several years, as in the Jena experiment. It is worth noting at this point that there is a mismatch of time scale between the modelled processes of plant growth occurring at a daily time step, and the available biomass data collected at a seasonal time step, at the end of each growing season. This mismatch of time scales requires developing an adapted estimation method detailed below.

The sampling procedure used to collect biomass data generates an observation error. We model this observation error with a lognormal distribution, meaning that observed biomass

$B_i^{obs}(T,p,s)$  of species  $i$  in season  $T$ , plot  $p$ , sample  $s$ , is the real biomass  $B_i(T,p)$  in season  $T$ ,

202 plot  $p$ , plus a random variable representing observation error:

$$\ln(B_i^{obs}(T, p, s)) = \ln(B_i(T, p)) + \sigma_{oi} \cdot u_{oi}(T, p, s) \quad (\text{eq.2})$$

204 where  $\sigma_{oi}^2$  is the observation variance, and  $u_{oi}(T,p,s)$  are independent normal variables with zero mean and unit variance. The real biomass is therefore estimated as the average of

206 observed log-transformed biomass across samples:

$$\hat{B}_i(T, p) = \exp\left(\frac{1}{N_s} \sum_{s=1}^{N_s} \ln(B_i^{obs}(T, p, s))\right) \quad (\text{eq.3})$$

208 where  $N_s$  is the number of samples harvested in a plot. Note that we used here the notation

$\hat{B}_i(T, p)$  to designate our estimates of the true values  $B_i(T,p)$ . We keep this convention for

210 estimates in the following of this paragraph. The full model including the dynamical model (eq. 1) and the observation process (eq. 2) therefore contains five parameters per species ( $r_{mi}$ ,

212  $K_i$ ,  $\sigma_{ei}$ ,  $\sigma_{di}$ ,  $\sigma_{oi}$ ) when applied to monoculture data.

214 The observation variance is estimated as the empirical variance of observed biomass across samples:

$$\hat{\sigma}_{oi}^2 = \frac{1}{N_T \cdot N_p \cdot (N_s - 1)} \cdot \sum_{T,p,s} \left( \ln(B_i^{obs}(T, p, s)) - \ln(\hat{B}_i(T, p)) \right)^2 \quad (\text{eq.4})$$

The estimated real biomass data  $\hat{B}_i(T, p)$  according to equation 3 (eq.3) are then used in the

218 following. It is important to note that our estimates of observation variance are based on the measured biomass variability among different samples of a same plot. Consequently, this

220 intra-plot variability is likely to be partly due to demographic variability at the sample scale.

Following de Mazancourt et al. (2013), we made the pragmatic choice to neglect this

222 demographic variability at the sample scale, since we do not have ways to disentangle the respective roles of observation and demographic noises at this scale. Our estimates of

demographic variability described below are thus to be understood as estimates of demographic variability *at the plot scale*.

The carrying capacity  $K_i$  of species  $i$  is assumed to be linearly related to the maximal biomass of species  $i$  across seasons, and is therefore estimated as:

$$\hat{K}_i = a \cdot \max_{T,p} \hat{B}_i(T, p) \quad (\text{eq.5})$$

In the following,  $a$  will be arbitrarily fixed to 2. We checked that alternative choices for  $a$  did not change the results qualitatively by repeating the analyses with values of  $a$  equal to 1.5 and 3 (data not shown).

We then estimated the intra-seasonal growth rates  $r_i(T, p)$  of species  $i$  for plot  $p$  and season  $T$ , defined as:

$$r_i(T, p) = r_{mi} + \sigma_{ei} \cdot u_{ei}(T) + \frac{\sigma_{di} \cdot u_{di}(T, p)}{\sqrt{B_0}} \quad (\text{eq.6})$$

For a given value of  $r_i(T, p)$ , the biomass dynamics can be estimated with the following difference equation (see eq. 1):

$$B_i^{sim}(t+1, p, T) = B_i^{sim}(t, p, T) + r_i(T, p) \cdot B_i^{sim}(t, p, T) \cdot \left( 1 - \frac{B_i^{sim}(t, p, T)}{\hat{K}_i} \right) \quad (\text{eq.7})$$

We used an optimization algorithm to estimate  $\hat{r}_i(T, p)$  that minimizes the difference between the simulated and real biomass at the end of the growing season  $\left| B_i^{sim}(60, p, T) - \hat{B}_i(T, p) \right|$ .

Then the set of model parameters ( $r_{mi}$ ,  $\sigma_{ei}$ ,  $\sigma_{di}$ ) are estimated from the estimated intra-seasonal growth rates  $\hat{r}_i(T, p)$  with a maximum likelihood procedure: let  $\hat{r}$  be the vector of previously

244 estimated intra-seasonal growth rates then the likelihood function  $L(\{r_{mi}, \sigma_{ei}, \sigma_{di}\} | \hat{r})$  can be  
 246 defined up to a multiplicative constant as:

$$L(\{r_{mi}, \sigma_{ei}, \sigma_{di}\} | \hat{r}) = \prod_T \int_{-\infty}^{+\infty} \frac{1}{(\sqrt{2\pi})^{Np+1} \sigma_{ei}^{Np} \sigma_{di}^{Np}} \sqrt{B_0}^{Np} \exp \left[ -\frac{1}{2} \cdot \left( \frac{t}{\sigma_{ei}} \right)^2 + \sum_p \left( \frac{\hat{r}_i(T, p) - r_{mi} - t \sqrt{B_0}}{\sigma_{di}} \right)^2 \right] dt \quad (\text{eq.8})$$

248 This equation (eq. 8) is obtained by writing down the probability of generating the estimated  
 intra-seasonal growth rates  $\hat{r}_i(T, p)$  from random draws of  $u_{ei}(T)$  and  $u_{di}(T, p)$  in normal  
 250 distributions with zero mean and unit variance.

The set of model parameters ( $r_{mi}, \sigma_{ei}, \sigma_{di}$ ) are estimated as:

$$252 \quad \left\{ \hat{r}_{mi}, \hat{\sigma}_{ei}, \hat{\sigma}_{di} \right\} = \arg \max L(\{r_{mi}, \sigma_{ei}, \sigma_{di}\} | \hat{r}) \quad (\text{eq.9})$$

Finally the estimation of environmental and demographic variables  $u_{ei}(T)$  and  $u_{di}(T, p)$  is based

254 on a maximum likelihood procedure. Let  $\hat{r}(T, \cdot) = \left( \hat{r}(T, 1), \dots, \hat{r}(T, p) \right)$  be the vector of previously  
 estimated intra-seasonal growth rates across plots for season  $T$  and  $\mathbf{u}(T) = (u_{ei}(T), \mathbf{u}_{di}(T, \cdot))$   
 256 with  $\mathbf{u}_{di}(T, \cdot)$  the vector of demographic random draws across plots for season  $T$ . From (eq.6),  
 the likelihood  $L(u(T) | \hat{r}(T, \cdot), \hat{r}_{mi}, \hat{\sigma}_{ei}, \hat{\sigma}_{di})$  is given by:

$$258 \quad L(u(T) | \hat{r}(T, \cdot), \hat{r}_{mi}, \hat{\sigma}_{ei}, \hat{\sigma}_{di}) = \frac{1}{(\sqrt{2\pi})^{Np+1} \hat{\sigma}_{ei} \hat{\sigma}_{di}^{Np}} \sqrt{B_0}^{Np} \cdot \exp \left[ -\frac{1}{2} \cdot \left( u_{ei}(T)^2 + \sum_p \left( \frac{\hat{r}_i(T, p) - \hat{r}_{mi} - \hat{\sigma}_{ei} \cdot u_{ei}(T)}{\hat{\sigma}_{di}} \right)^2 \right) \right] \quad (\text{eq.10})$$

260 Maximisation of the likelihood function enables to estimate  $\hat{u}_{ei}(T)$ . The demographic values  
 $\hat{u}_{di}(T, \cdot)$  are then deduced from (eq.6). All optimization procedures were performed with the

“optim” R function using the Nelder and Mead (1965) algorithm. In the following, we do not distinguish the notation of estimated values from that of true values to keep the notations simple.

#### 2.4. Model checking

To assess the ability of our model to reproduce observed biomass dynamics in monocultures, we performed several model checks. First, we assessed, for each of the 60 species in monocultures, whether fitted values of environmental variables  $u_{ei}(T)$ , demographic variables  $u_{di}(T,p)$  and observation errors  $u_{oi}(T,p,s)$  significantly deviated from normal distributions using a Kolmogorov-Smirnov test. Second, we simulated the monospecific biomass dynamics using our stochastic model with fitted parameters  $(r_{mi}, K_i, \sigma_{ei}, \sigma_{di}, \sigma_{oi})$  and variables  $(u_{ei}, u_{di}, u_{oi})$  randomly drawn in normal distribution with zero mean and unit variance and compared the outputs of these simulations with observed biomass dynamics to assess whether systematic deviations between the two were present. More precisely, for each of the 60 species, we simulated virtual data, following the Jena experimental setup: each monoculture was replicated in two plots, and two subsamples by plot were harvested during 17 seasons. Each such monoculture simulation was replicated 100 times, and for each replication, the average and standard deviation of community biomass across plots, subsamples and seasons were computed. The ability of the model to reproduce the observed values of these indicators for species  $i$  was assessed through the p-value  $p_i$  computed as the rank of the observed value within the set of simulated values. The model was considered to pass the model checking test when the p-value was comprised between 0.025 and 0.975.

#### 2.5. Assessing the need for the current level of model complexity

We also considered three degraded versions of the model, in order to demonstrate that the current level of model complexity is necessary to satisfactorily model the empirical observations in monocultures. The first alternative model neglects demographic variance, and assumes that  $\sigma_{di} = 0$  for all species  $i$ . The second alternative model further neglects environmental variance and assumes that  $\sigma_{ei} = 0$  for all species  $i$ . The third alternative model further assumes that the three remaining model parameters ( $r_m, K, \sigma_o$ ) are constant across species. Likelihood formula for these alternative models were derived (Table S1) and model selection was performed via AIC computations (Table S2). These analyses confirmed that the full model clearly outperformed these degraded versions based on the monoculture data (Table S2), so that we do not longer consider these alternative models in the remaining analyses.

## 2.6. Quantifying the sources of biomass variability

For each species, we assessed the effects of environmental stochasticity, demographic stochasticity and observation error on observed biomass variances. To do this, we simulated monoculture biomass dynamics using fitted parameters ( $r_{mi}, K_i, \sigma_{ei}, \sigma_{di}, \sigma_{oi}$ ) and variables ( $u_{ei}, u_{di}, u_{oi}$ ) randomly drawn in normal distribution with zero mean and unit variance in a virtual plot sampled once during 1000 seasons. This biomass time series was used to compute the reference biomass variance  $V_{ref}$ . We then replicated this procedure, shutting down sequentially observation error, which generates variability between samples (and plots and years), and demographic stochasticity, which generates variability between plots (and years),. The contributions of the three sources of variability to observed biomass variability were then computed as the relative decrease in biomass variance caused by the suppression of these sources of variability in the simulations. More precisely, the effect of observation error was computed as  $\Sigma_o = (V_{ref} - V_{e+d}) / V_{ref}$ , where  $V_{e+d}$  stands for the biomass variance in

simulations with environmental and demographic stochasticity only. The effect of  
 312 demographic stochasticity was computed as  $\Sigma_d = (V_{e+d} - V_e) / V_{ref}$ , where  $V_e$  stands for the  
 biomass variance in simulations with environmental stochasticity only. The effect of  
 314 environmental stochasticity was finally computed as  $\Sigma_e = V_e / V_{ref}$ . This asymmetry in our  
 way to compute these three variance components stems from the assumed nested structure of  
 316 these three sources of variability.

### 318 *2.7. Analyzing species demographical characteristics*

We assessed whether species with larger intrinsic growth rates  $r_{mi}$  have also larger  
 320 demographic  $\sigma_{di}^2$  and environmental  $\sigma_{ei}^2$  variances by correlation analyses. We also assessed  
 whether species demographical characteristics ( $r_{mi}$ ,  $K_i$ ,  $\sigma_{ei}$ ,  $\sigma_{di}$ ) could be predicted from three  
 322 functional traits: plant vegetative height, specific leaf area (SLA) and leaf dry matter content  
 (LDMC). Trait data were extracted from the LEDA trait database (Kleyer et al., 2008). When  
 324 multiple entries for a given species were available, we used average values.

### 326 *2.8. Predicting multi-specific community dynamics based on the assumption of*

#### 328 *competitive symmetry*

We finally assessed whether knowledge of species individual demographical characteristics  
 330 was sufficient to predict the multi-specific community dynamics in mixtures. Predictions in  
 mixtures were based on the assumption of competitive symmetry, meaning that the  
 332 coefficients  $\alpha_{ij}$  in equation 1 were fixed to 1. This assumption is rather strong, but it represents  
 a first step in order to predict complex community dynamics, in the spirit of neutral  
 334 community models (Rosindell et al. 2011) and their extensions (Jabot and Chave 2011,

Kalyuzhny et al. 2015). All other species demographical parameters ( $r_{mi}$ ,  $K_i$ ,  $\sigma_{ei}$ ,  $\sigma_{di}$ ,  $\sigma_{oi}$ ) were  
 336 set to their estimates from monoculture data (see section 2.3).

338 We performed several sets of predictions for mixtures to assess the gain in prediction  
 accuracy brought by several components of species demography. In the first set of  
 340 simulations, we used the values of environmental variables  $\hat{u}_{ei}(T)$  estimated from monoculture  
 data. The demographic and observation variables,  $u_{di}(T,p)$  and  $u_{oi}(T,p,s)$  are however likely to  
 342 vary from one plot to another. We therefore run 100 simulations for each mixture with  
 $u_{di}(T,p)$  and  $u_{oi}(T,p,s)$  randomly drawn in normal distributions with zero mean and unit  
 344 variance, and using the same sampling protocol as in the Jena experiment (see section 2.2).  
 This set of simulation incorporates the observed temporal synchrony or asynchrony between  
 346 species in their responses to environmental variability, since we used the empirically  
 estimated values of environmental variables  $\hat{u}_{ei}(T)$ . We additionally performed a second set  
 348 of simulations using environmental variables  $u_{ei}(T)$  randomly drawn in normal distributions  
 with zero mean and unit variance. In this second set of simulations, species are therefore  
 350 assumed to respond independently to environmental variability, and the observed temporal  
 synchrony/asynchrony between species is randomized.

352  
 For each set of simulations, we computed the average  $\bar{B}$  and standard deviation  $\sigma_B$  of  
 354 community biomass across seasons and subsamples, as well as the Simpson's concentration  
 index defined as:

$$\lambda = \sum_{i=1}^N \overline{p_{bi}}^{-2} \quad (\text{eq.11})$$

Where  $\overline{p_{bi}}$  is the average proportion of community biomass attributable to species  $i$ .



For each mixture, the distribution across replicates of  $\bar{B}$ ,  $\sigma_B$  and  $\lambda$  was compared with  
 360 observed values and the ability of the model to reproduce community dynamics was assessed  
 through the p-value computed as the rank of the observed value within the set of simulated  
 362 values. For each species richness level, we defined an aggregated indicator  $N_{\text{dev}}$ , which  
 describes the proportion of mixtures for which the p-value is smaller than 0.025 or higher than  
 364 0.975. Moreover the average of  $\bar{B}$ ,  $\sigma_B$  and  $\lambda$  across replicates were computed and the  
 correlation between predicted averages and observed values across mixtures was quantified.

366  
 Finally, the overall discrepancy between observed values in mixtures and predicted averages  
 368 is not necessarily indicative of a failure of the assumption of competitive symmetry. Indeed, it  
 may be due solely to demographic stochasticity and observation errors, which are averaged  
 370 out in predicted averages but not in observed values. We therefore computed the correlation  
 coefficient across mixtures between the predicted average and one simulation value randomly  
 372 picked among the 100 replicates for each mixture,  $R^2_{\text{samp}}$ . This procedure is repeated 100  
 times and the validity of the assumption of competitive symmetry is assessed through the p-  
 374 value, computed as the rank of the correlation coefficient between observed and predicted  
 values within the set of  $R^2_{\text{samp}}$  obtained through the sampling procedure. The assumption of  
 376 competitive symmetry is rejected when the p-value is smaller than 0.025 or larger than 0.975.

378 All R scripts are available on Github (<https://github.com/thlohier/SGCDM>).

### 380 **3. Results**

#### *3.1. Model checking and inference based on monoculture data*

382 Our model passed most model checking tests successfully. Indeed, only two out of the 60  
 species had not normally distributed environmental variables  $u_{ei}$  according to the

Kolmogorov-Smirnov test (Figure 2A): *Geranium pratense* and *Luzula campestris*. And only one species had not normally distributed demographic variables  $u_{di}$  according to the Kolmogorov-Smirnov test (Fig.2.B): *Trifolium fragiferum*. Model failure for *G. pratense* was likely due to the relatively large between plot variability that lead to a shrinked distribution of environmental variables. Model failure for *L. campestris* and *T. fragiferum* can be explained by the fact that they were frequently not found in the plots: *L. campestris* was not found for 8 of the 14 seasons and *T. fragiferum* was found in only one of the two plots in which it was sown. Moreover *C. pratensis* was not found for 13 of the 14 seasons. These four species were thus excluded from the subsequent analyses. Model failure was also detected for three additional species on other model checks (Figure 2, Table S4). These species were however kept in the subsequent analyses because model failure was more marginal in these cases. For the remaining 56 species, we were able to estimate the four species demographical parameters ( $r_{mi}$ ,  $K_i$ ,  $\sigma_{ei}$ ,  $\sigma_{di}$ ) (Table 1).

### 3.2. Sources of biomass variability

Our analysis revealed that the variance in community biomass  $\sigma_B^2$  was mainly due to observation error with an average proportion of explained variance equal to 44%, followed by environmental stochasticity (34%) and demographic stochasticity (22%). More precisely, observation error was the main cause of biomass variance for 30 out of the 56 species, while environmental stochasticity was the main cause for 22 out of the 56 species (Table 1). We further detected variations in this hierarchy depending on the functional group considered, with the effect of environmental stochasticity being especially large for grasses, especially low for legumes, and intermediate for tall and small herbs. In contrast, the effect of observation error was more consistently large, while the one of demographic stochasticity more consistently low across the different functional groups (Figure 3).

### 3.3. Analysis of species demographical characteristics

Overall, we did not find any support for the idea of an equalizing trade-off in the plant demographical components. Intrinsic growth rates ( $r_m$ ) tended to be negatively correlated with environmental ( $\sigma_e^2$ ), demographic ( $\sigma_d^2$ ) and observation ( $\sigma_o^2$ ) variance while observation variance tended to be positively correlated with environmental and demographic variance (Table S5). The correlations were weak when the functional types were not distinguished ( $R^2$  close to 0.3). But we observed a significant negative correlation between  $r_m$  and  $\sigma_e$  as well as between  $r_m$  and  $\sigma_o$  for grass species ( $R^2 = 0.62$  and  $R^2=0.40$  respectively). There was also a significant negative correlation between  $r_m$  and  $\sigma_o$  for small herbs, between  $r_m$  and  $\sigma_e$  for legumes and between  $\sigma_d$  and  $\sigma_o$  for grasses and tall herbs ( $R^2 > 0.4$ ).

We found that some plant demographical characteristics ( $r_m$  and  $K$ ) were moderately predictable from the functional traits examined. When all plant functional groups were considered, both intrinsic growth rate  $r_m$  and carrying capacity  $K$  were positively correlated with plant height and LDMC, and negatively with plant SLA (Table 2). Four of these correlations were significant ( $r_m - \text{Height}$ ,  $r_m - \text{SLA}$ ,  $K - \text{SLA}$ ,  $K - \text{LDMC}$ ), and one marginally significant ( $K - \text{Height}$ ). However, depending on the functional group considered, the link between plant functional traits and these demographical parameters varied: plant height and LDMC were mostly influential for grasses, while SLA mostly influenced herbs and legumes. In contrast, the plant functional traits examined had little explanatory power on both environmental and demographic variances, neither on observation errors (Table 2).

### 3.4. Predicting multi-specific community dynamics based on the assumption of competitive symmetry

We obtained a good predictive ability of multi-species community dynamics under the assumption of competitive symmetry (Figure 4), with coefficient of determinations between predicted and observed values equal to 0.65 and 0.60 for biomass averages and standard deviations, and to 0.75 for community composition (Simpson's concentration index). We further found that our predictive ability tended to decrease when considering mixtures of increasing diversity (Figure 4, Table 3). Surprisingly, we found that taking into account the observed asynchrony between species in their response to environmental variations did not improve our predictive ability of these summary statistics (Supplementary Table S6).

In these comparisons between predicted and observed community dynamics, predictions are computed from averages over 100 simulated community dynamics, while observed values are based on single realizations of this dynamics and single realizations of observation errors. We therefore assessed whether the discrepancy between observations and predictions was due to model error or whether it may only be due to demographic stochasticity and observation error using two complementary tests. First, we assessed for each mixture, whether observed values fell within the distribution of the 100 simulated values. We found that only about one fifth of mixtures significantly deviated from the simulated distributions ( $N_{dev}$  in Table 3). Second, to obtain a more synthetic picture of the magnitude of model error, we assessed whether the coefficient of determination between observed and predicted values across communities significantly deviated from the null distribution based on simulated communities. Using this approach, our model was not rejected regarding its predictions on biomass averages and standard deviations, but it was rejected regarding its predictions on community composition (Simpson's concentration index) that overestimated the species abundance evenness of the communities.

## DISCUSSION

The analysis of stochastic community dynamics has been mainly carried on easily countable organisms (Lande et al. 2003, Chisholm et al. 2014, Kalyuzhny et al. 2015). The present study aimed at extending such analyses to herbaceous plants which are easier to monitor in terms of aerial biomass, which experience inter-individual competition at short time scales, and which suffer from sudden seasonal collapses due to winter mortality or agricultural management. We therefore proposed a novel model of stochastic community dynamics, inspired by the recent work of de Mazancourt et al. (2013). We further developed an original inference method based on both numerical simulations and likelihood computing so as to tackle the mismatch of time scales between modelled processes occurring at a daily time step and field observations gathered at a seasonal time step (Figure 1). This inference method further has the advantage of being very flexible, so that alternative modelling assumptions can be easily explored within the same quantitative framework.

Our simple model was sufficient to reproduce most observed monoculture data of the Jena biodiversity experiment (Figure 2). We further found that we could predict multi-specific community dynamics with a relatively good precision using species demographical parameters estimated from monoculture data and assuming competitive symmetry, the fact that intraspecific and interspecific competitions have the same magnitude (Figure 4, Table 3). Still, further improvements of our predictive ability should be sought in relaxing this simplifying assumption of competitive symmetry. In particular, the fact that our model predictions overestimated the species abundance evenness of the communities indicates that some level of asymmetry cannot be ignored to fully understand the dynamics of these herbaceous communities. This result is in accord with previous knowledge on plant

competitive asymmetry (Begon, 1984, Cannell et al., 1984). It further provides clues on how to proceed to improve the model in the future. For instance, competitive asymmetry between a pair of plant species might be predicted from their difference in functional trait values, so that plant functional traits may be included in the model, as drivers of competition coefficients.

Our approach further revealed that environmental stochasticity had a larger impact on population temporal variance than demographic stochasticity (Figure 3). The importance of environmental stochasticity on community dynamics has recently received a renewed interest (Chisholm et al. 2014, Kalyuzhny et al. 2015). Our results are perfectly in line with these findings in other types of plant communities. While environmental variability is likely to foster species coexistence by favoring in turn different sets of species (Chesson and Warner 1981), our analysis revealed that it was unlikely to be involved in an equalizing trade-off among species between their intrinsic growth rate and their temporal stability. Indeed, we found the contrary result that species with larger growth rates also tended to have both lower demographic variance and lower environmental variance (Table S5).

We finally found that the plant functional traits examined only had modest predictive ability of the species demographical characteristics (Table 2). This result recalls the recent findings of Kraft et al. (2015) on pairwise competition between annual plants. As previously argued, plant functional traits may however be informative on the competitive strength between pairs of species. This clearly constitutes a promising perspective to improve the present state of modelling of the stochastic dynamics of herbaceous plant communities.

More generally, among the growing literature on stochastic community dynamics (Freckleton et al. 2000; Keith et al. 2008; Loreau and de Mazancourt 2008; Chase et al. 2010; Fowler et

*al.* 2012), our approach is original in that it jointly uses simulation-based and analytical  
inference methods to study more complex models of stochastic community dynamics than  
those commonly used. By using more detailed and realistic representations of biological  
processes, such hybrid inference approaches should contribute to the development of complex  
models of ecosystem dynamics (Evans *et al.* 2013) that can be informed by field data.

## ACKNOWLEDGEMENTS

T.L was funded by the Institut national de Recherche en Sciences et Technologie pour  
l'Environnement et l'Agriculture (IRSTEA) and the Regional Council of Auvergne. We thank  
the Editor and an anonymous reviewer for their constructive comments that improved the  
manuscript.

## REFERENCES

- 524 Ackerly, D.D., 2003. Community assembly, niche conservatism, and adaptive evolution in  
changing environments. *Int J. Plant Sci.* 164: S165-S184.
- 526 Beaumont, M.A., 2010. Approximate Bayesian computation in population genetics. *Annu.*  
*Rev. Ecol. Evol. Syst.* 41: 379-406.
- 528 Begon, M., 1984. Density and individual fitness: asymmetric competition. In: Shorrocks, B.  
(ed.). *Evolutionary ecology*. Blackwell, Oxford.
- 530 Cannell, M.G.R., et al., 1984. Competition within stand of *Picea sitchensis* and *Pinus*  
*contorta*. *Ann. Bot.* 53: 349-362.
- 532 Chase, J.M., 2010. Stochastic community assembly causes higher biodiversity in more  
productive environments. *Science* 328: 1388-1391.
- 534 Chesson, P.L., Warner, R.R., 1981. Variability promotes coexistence in lottery competitive  
systems. *Am. Nat.* 117: 923-943.
- 536 Chesson, P.L., 2000. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Evol.*  
*Syst.* 31: 343-538.
- 538 Chisholm, R.A., et al., 2014. Temporal variability of forest communities: empirical estimates  
of population change in 4000 tree species. *Ecol. Lett.* 17: 855-865.
- 540 Cornwell, W.K., et al., 2006. A trait-based test for habitat filtering: convex Hull volume.  
*Ecology* 87: 1465-1471.
- 542 De Mazancourt, C., et al., 2013. Predicting ecosystem stability from community composition  
and biodiversity. *Ecol. Lett.* 16: 617-625.
- 544 Ejrnaes, R., et al., 2006. Community assembly in experimental grasslands: suitable  
environment or timely arrival. *Ecology* 87: 1225-1233.
- 546 Evans, M. R., et al., 2013. Do simple models lead to generality in ecology? *Trends Ecol.*  
*Evol.* 28: 578-583.



548 Fowler, M. S., et al., 2012. Species dynamics alter community diversity–biomass stability  
relationships. *Ecol. Lett.* 15: 1387-1396.

550 Freckleton, R.P., et al., 2000. Determinants of the abundance of invasive annual weeds:  
community structure and non-equilibrium dynamics. *Proc. Roy. Soc. B.* 267: 1153-1161.

552 Gonzalez, M.A., et al., 2010. Shifts in species and phylogenetic diversity between sapling and  
tree communities indicate negative density dependence in a lowland rain forest. *J. Ecol.* 98:  
554 137-146.

Hector, A., et al., 2010. General stabilizing effects of plant diversity on grassland productivity  
556 through population asynchrony and overyielding. *Ecology* 91: 2213-2220.

Jabot, F., Chave, J., 2011. Analyzing tropical forest tree species abundance distributions using  
558 a nonneutral model and through approximate Bayesian inference. *Am. Nat.* 178: E37-E47.

Jabot, F., Pottier, J., 2012. A general modelling framework for resource-ratio and CSR  
560 theories of plant community dynamics. *J. Ecol.* 100: 1296-1302.

Jouven, M., et al., 2006. Model predicting dynamics of biomass, structure and digestibility of  
562 herbage in managed permanent pastures. 2. Model evaluation. *Grass For. Sci.* 61: 125-133.

Kalyuzhny, M., et al., 2015. A neutral theory with environmental stochasticity explains static  
564 and dynamic properties of ecological communities. *Ecol. Lett.* 18: 572-580.

Keith, D. A., et al., 2008. Predicting extinction risks under climate change: coupling  
566 stochastic population models with dynamic bioclimatic habitat models. *Biol. Lett.* 4: 560-  
563.

568 Kraft, N.J., et al., 2015. Plant functional traits and the multidimensional nature of species  
coexistence. *Proc. Natl. Acad. Sci. USA* 112: 797-802.

570 Lande, R., et al., 2003. Stochastic population dynamics in ecology and conservation. Oxford  
University Press, New York.

572 Loreau, M., de Mazancourt, C., 2008. Species synchrony and its driver: neutral and  
 nonneutral community dynamics in fluctuating environments. *Am. Nat.* 172: E48-E66.

574 Maestre, F. T., et al. 2009. Refining the stress-gradient hypothesis for competition and  
 facilitation in plant communities. *J. Ecol.* 97: 199-205.

576 Mouquet, N., et al., 2002. Plant species richness and community productivity: why the  
 mechanism that promotes coexistence matters. *Ecol. Lett.* 5: 56-65.

578 Mouquet, N. et al., 2015. Predictive ecology in a changing world. *J. App. Ecol.* 52: 1293-  
 1310.

580 Nelder, J. A., Mead, R., 1965. A simplex algorithm for function minimization. *Comp. J.* 7:  
 308-313.

582 Norden, N., et al., 2009. Resilience of tropical rain forests: tree community reassembly in  
 secondary forests. *Ecol. Lett.* 12: 385-394.

584 Raven, P. H., Evert, R. F., & Eichhorn, S. E. (2005). *Biology of plants*. Macmillan.

Roscher, C., et al., 2004. The role of biodiversity for element cycling and trophic interactions:  
 586 an experimental approach in a grassland community. *Bas. Appl. Ecol.* 5: 107-121.

Rosindell, J., et al., 2011. The unified neutral theory of biodiversity and biogeography at age  
 588 ten. *Trends Ecol. Evol.* 26: 340-348.

Silvertown, J., et al., 2015. Hydrological niches in terrestrial plant communities: a review. *J.*  
 590 *Ecol.* 103: 93-108.

Sprugel, D.G. 1991. Disturbance, equilibrium, and environmental variability: what is 'natural'  
 592 vegetation in a changing environment? *Biol. Cons.* 58: 1-18.

Stubbs, W.J., Wilson, J.B., 2004. Evidence for limiting similarity in a sand dune community.  
 594 *J. Ecol.* 92: 557-567.

Weigelt, A., et al., 2010. The Jena experiment: six years of data from a grassland biodiversity  
 596 experiment. *Ecology* 91: 930.

Wilson, S. D., Tilman, D., 1993. Plant competition and resource availability in response to

598 disturbance and fertilization. *Ecology* 74: 599-611.

600

## TABLES

**Table 1.** Species demographical characteristics estimated from monoculture data including the intrinsic growth rate ( $r_m$ ), the carrying capacity ( $K$ ), the environmental ( $\sigma_e$ ), demographic ( $\sigma_d$ ) and observation ( $\sigma_o$ ) standard deviations as well as the relative effects of environmental ( $\Sigma_e$ ) and demographic ( $\Sigma_d$ ) stochasticities and observation error ( $\Sigma_o$ ) on community biomass variance.

Species	$r_m$	$K$	$\sigma_e$	$\sigma_d$	$\sigma_o$	$\Sigma_e$	$\Sigma_d$	$\Sigma_o$
<b>Grasses</b>								
Alo.pra	0.074	1011.494	0.018	0.008	0.228	0.6	0.2	0.19
Ant.odo	0.06	668.109	0.025	0.015	0.463	0.35	0.17	0.48
Arr.ela	0.083	1510.984	0.013	0.008	0.279	0.46	0.24	0.3
Ave.pub	0.067	1267.008	0.022	0.006	0.432	0.52	0.08	0.41
Bro.ere	0.081	2039.401	0.015	0.009	0.379	0.37	0.21	0.41
Bro.hor	0.057	964.486	0.024	0.02	0.834	0.07	0.09	0.84
Cyn.cri	0.06	1134.767	0.026	0.011	0.488	0.36	0.2	0.44
Dac.glo	0.038	289.6	0.022	0.012	0.443	0.64	0.18	0.17
Fes.pra	0.059	1502.602	0.026	0.014	0.445	0.26	0.35	0.39
Fes.rub	0.064	1250.319	0.018	0.015	0.38	0.56	0.33	0.11
Hol.lan	0.04	367.476	0.026	0.011	0.588	0.53	0.23	0.24
Phl.pra	0.092	1269.795	0.009	0.008	0.198	0.25	0.24	0.51
Poa.pra	0.067	714.514	0.012	0.008	0.295	0.22	0.12	0.66
Poa.tri	0.063	510.719	0.011	0.021	0.435	0.22	0.14	0.64
Tri.flu	0.05	837.552	0.019	0.014	0.546	0.82	0.06	0.12
<b>Small herbs</b>								
Aju.rep	0.029	68.62	0.009	0.016	0.476	0.08	0.33	0.58
Bel.per	0.023	162.489	0.02	0.011	0.926	0.07	0.05	0.88
Gle.hed	0.076	940	0.01	0.008	0.262	0.29	0.11	0.6
Leo.aut	0.083	1549.742	0.014	0.005	0.234	0.57	0.2	0.22
Leo.his	0.068	718	0.017	0.007	0.296	0.47	0.32	0.22
Pla.lan	0.059	459.972	0.018	0.007	0.437	0.08	0.42	0.5
Pla.med	0.067	1383.776	0.021	0.016	0.484	0.7	0.07	0.23
Pri.ver	0.074	869.573	0.019	0.005	0.266	0.02	0.06	0.91
Pru.vul	0.055	652.037	0.019	0.01	0.611	0.24	0.18	0.58
Ran.rep	0.048	656.356	0.019	0.021	0.995	0.21	0.15	0.64
Tar.off	0.044	445.021	0.021	0.012	0.603	0.47	0.08	0.45
Ver.cha	0.078	880.468	0.02	0.003	0.439	0.55	0.14	0.31

Tall herbs								
Ach.mil	0.069	1229.282	0.017	0.005	0.148	0.76	0.15	0.09
Ant.syl	0.053	1113.886	0.025	0.016	0.565	0.2	0.21	0.59
Cam.pat	0.03	968.599	0.035	0.008	0.646	0.36	0.06	0.58
Car.car	0.046	721.216	0.013	0.013	0.353	0.19	0.44	0.37
Cen.jac	0.001	4.1	0	0	NA	0.67	0.17	0.15
Cir.ole	0.072	1506.494	0.018	0.007	0.212	0.4	0.28	0.32
Cre.bie	0.061	696.114	0.017	0.011	0.32	0.38	0.14	0.48
Dau.car	0.078	1225.749	0.015	0.006	0.202	0.35	0.22	0.43
Gal.mol	0.068	1129.696	0.02	0.012	0.162	0.35	0.33	0.32
Her.sph	0.053	705.614	0.002	0.024	0.683	0.06	0.1	0.84
Kna.arv	0.047	603.754	0.018	0.016	0.831	0.64	0.14	0.22
Leu.vul	0.059	692.424	0.022	0.01	0.271	0.58	0.27	0.15
Pas.sat	0.04	478.295	0.021	0.018	0.645	0.41	0.11	0.48
Pim.maj	0.099	2446.728	0.011	0.01	0.634	0.39	0.28	0.33
Ran.acr	0.037	499.6	0.028	0.015	0.628	0.55	0.22	0.24
Rum.ace	0.047	936.789	0.023	0.013	0.569	0.12	0.33	0.55
San.off	0.063	590.141	0.015	0.008	0.252	0.08	0.51	0.41
Tra.pra	0.058	749.989	0.015	0.019	0.493	0.19	0.23	0.58
Legumes								
Lat.pra	0.079	1186.744	0.014	0.008	0.243	0.55	0.16	0.29
Lot.cor	0.073	824.344	0.012	0.008	0.215	0.26	0.36	0.38
Med.lup	0.07	1558.09	0.022	0.011	0.206	0.11	0.2	0.68
Med.var	0.075	853.241	0.012	0.013	0.328	0.39	0.35	0.26
Ono.vic	0.014	42.213	0.002	0.018	0.315	0.11	0.11	0.79
Tri.cam	0.08	941.01	0.004	0.01	0.285	0.41	0.14	0.45
Tri.dub	0.068	547.478	0.016	0.005	0.397	0.14	0.51	0.34
Tri.hyb	0.023	176.49	0.024	0.009	0.471	0.06	0.05	0.88
Tri.pra	0.028	405.337	0.023	0.026	0.358	0.01	0.33	0.65
Tri.rep	0.07	1125.347	0.023	0.004	0.201	0.23	0.4	0.37
Vic.cra	0.047	959.132	0.028	0.017	0.949	0.05	0.36	0.59

608

\*Full species names corresponding to each code are given in Table S3.

**Table 2.** Correlations between demographical parameters including the intrinsic growth rate ( $r_m$ ), the carrying capacity ( $K$ ), the environmental ( $\sigma_e$ ), demographic ( $\sigma_d$ ) and observation ( $\sigma_o$ ) standard deviations and three plant functional traits: plant height, specific leaf area (SLA) and leaf dry matter content (LDMC). The correlations were computed either for all species together or separately for each functional group. Significance at the 0.05 level are indicated in bold.

		Height		SLA		LDMC	
		R	p-value	R	p-value	R	p-value
$r_m$	All	<b>0.28</b>	<b>0.044</b>	<b>-0.40</b>	<b>0.003</b>	0.23	0.102
	Grass	0.39	0.167	-0.14	0.616	0.34	0.217
	Small herb	0.31	0.361	-0.47	0.122	0.13	0.712
	Tall herb	-0.01	0.976	<b>-0.49</b>	<b>0.038</b>	-0.04	0.889
	Legumes	0.14	0.69	-0.54	0.11	0.09	0.814
$K$	All	0.26	0.063	<b>-0.45</b>	<b>0.00</b>	<b>0.31</b>	<b>0.023</b>
	Grass	0.36	0.204	-0.35	0.205	0.45	0.095
	Small herb	0.03	0.923	-0.46	0.133	0.05	0.888
	Tall herb	-0.16	0.534	<b>-0.48</b>	<b>0.042</b>	0.01	0.966
	Legumes	-0.08	0.818	<b>-0.76</b>	<b>0.011</b>	0.22	0.542
$\sigma_e$	All	-0.17	0.224	<b>0.27</b>	<b>0.047</b>	0.06	0.687
	Grass	-0.49	0.072	0.09	0.745	-0.17	0.554
	Small herb	0.01	0.977	0.17	0.592	0.28	0.409
	Tall herb	-0.04	0.861	0.4	0.104	-0.36	0.162
	Legumes	-0.37	0.295	0.4	0.251	0.13	0.728
$\sigma_d$	All	-0.18	0.186	0.08	0.572	-0.10	0.461
	Grass	-0.51	0.062	0.32	0.252	-0.17	0.552
	Small herb	-0.18	0.594	-0.27	0.392	0.17	0.627
	Tall herb	0.24	0.345	0.15	0.548	-0.06	0.829
	Legumes	-0.16	0.656	0.14	0.699	-0.54	0.111
$\sigma_o$	All	-0.19	0.168	0.17	0.225	-0.05	0.727
	Grass	-0.48	0.083	0.19	0.494	0.03	0.911
	Small herb	-0.37	0.266	-0.09	0.779	-0.01	0.985
	Tall herb	0.38	0.117	<b>0.48</b>	<b>0.045</b>	0.00	0.992
	Legumes	-0.16	0.649	-0.3	0.4	0.19	0.595

**Table 3.** Assessment of model predictive ability in species mixtures.  $R^2$  between observed

and predicted values, proportion of mixtures  $N_{\text{dev}}$  significantly deviating from model

predictions and overall deviation significance  $p$  from predictions for the studied summary

statistics. Results regarding the community biomass average ( $\bar{B}$ ) and standard deviation ( $\sigma_B$ )

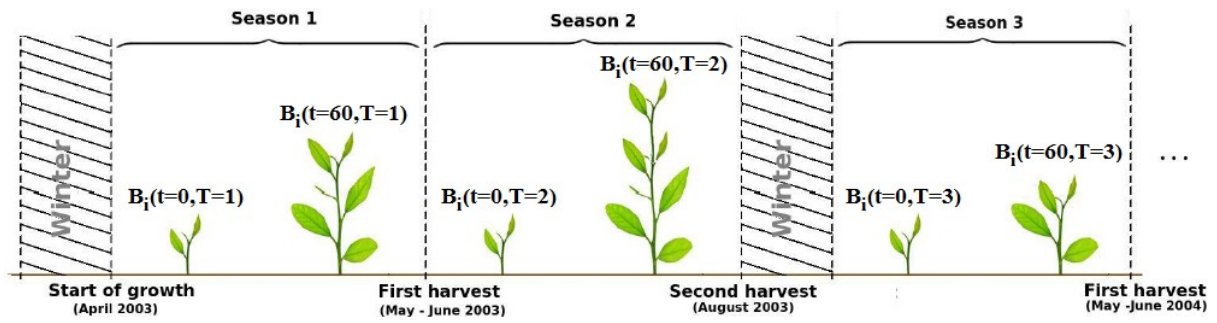
across seasons and subsamples as well as the Simpson's concentration index ( $\lambda$ ) are reported

both for all the empirical mixtures, and for each diversity level separately.

		All	2 species	4 species	8 species	16 species
$\bar{B}$	$R^2$	0.65	0.86	0.65	0.75	0.49
	$N_{\text{dev}}$	0.21	3/16	2/11	1/10	2/3
	$p$	0.02	0.48	0.33	0.43	0.45
$\sigma_B$	$R^2$	0.60	0.61	0.69	0.70	0.43
	$N_{\text{dev}}$	0.13	3/16	1/11	0/10	1/3
	$p$	0.31	0.17	0.20	0.49	0.49
$\lambda$	$R^2$	0.75	0.57	0.01	0.01	0.54
	$N_{\text{dev}}$	0.31	6/16	3/11	3/10	0/3
	$p$	0.00	0.01	0.00	0.04	0.37

## FIGURES

**Figure 1.** Model description and relationships with data used for calibration. Only end of season biomass  $B_i(t=60, T)$  is known.



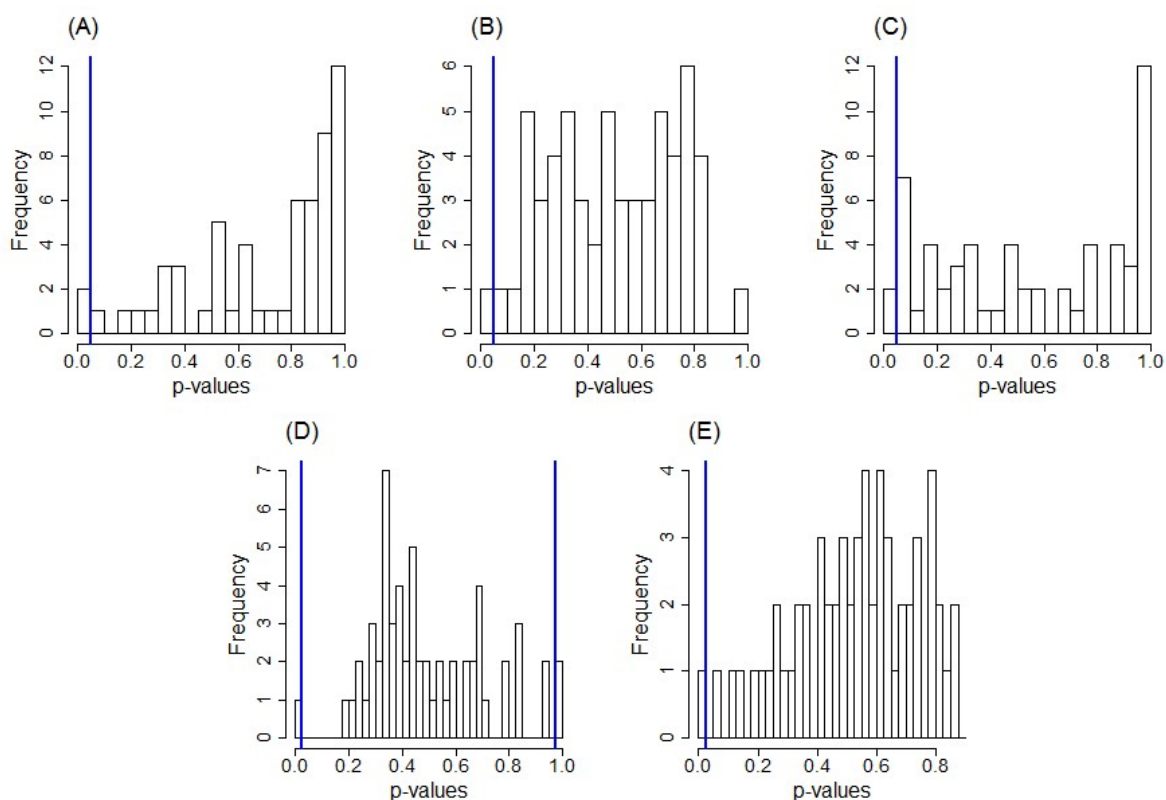


**Figure 2.** Model checking results. Normality test p-values of environmental variables  $u_{ei}$  (A),

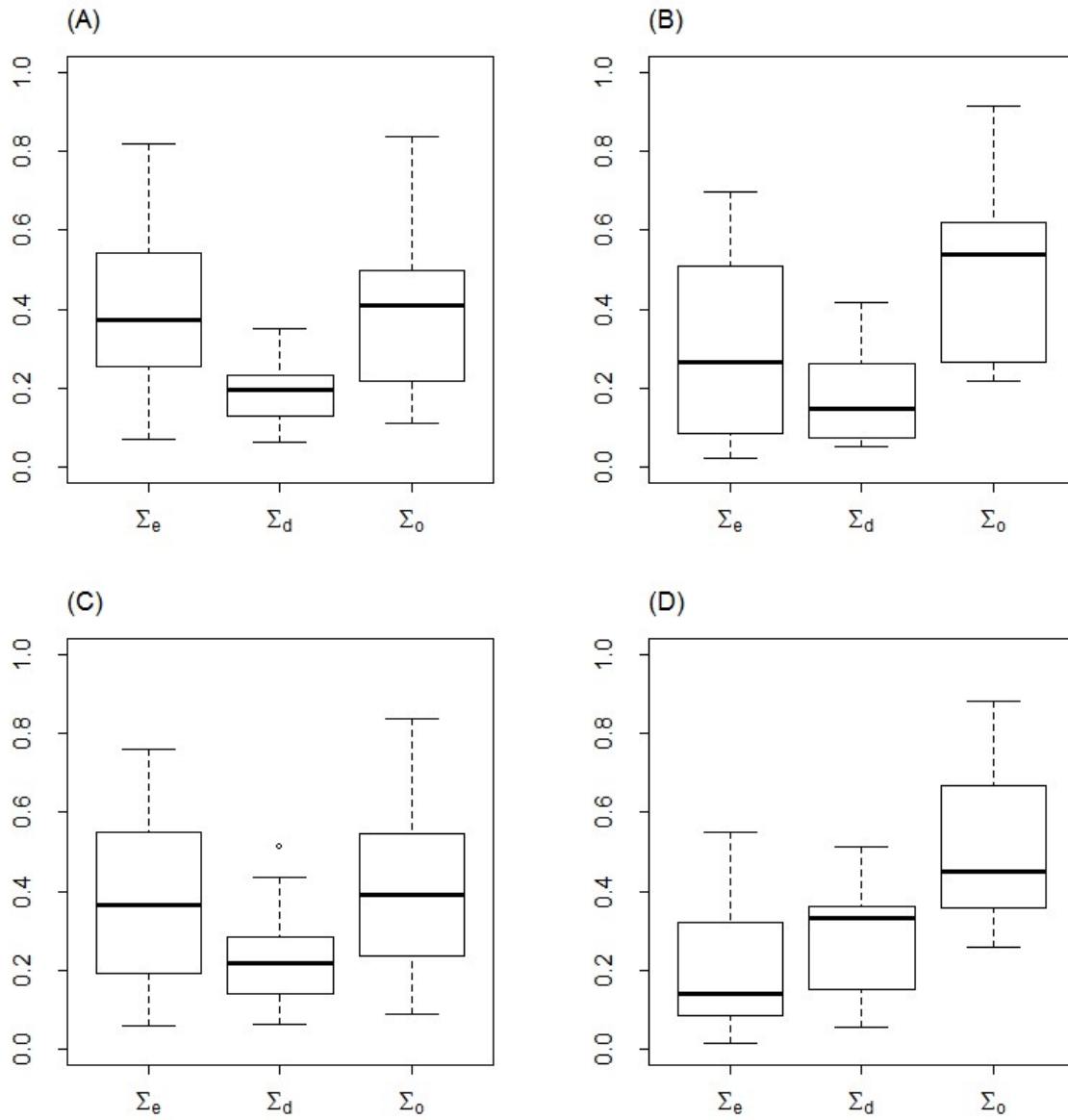
demographic variables  $u_{di}$  (B) and observation variables  $u_{oi}$  (C) obtained for each species.

And model checking p-values based on biomass averages (D) and standard deviations (E)

obtained for each species. The solid blue lines represent the one-tail 0.05 threshold for the first three panels, and the two-tail 0.025 and 0.975 thresholds for the last two panels.



642 **Figure 3.** Effects of observation error ( $\Sigma_o$ ), environmental ( $\Sigma_e$ ) and demographic ( $\Sigma_d$ )  
644 stochasticities on the variance of community biomass  $\sigma_B^2$  for (A) grasses; (B) small herbs; (C)  
644 tall herbs; (D) legumes.



646 **Figure 4.** Comparison between predicted and observed multi-species community dynamics,  
 based on their log-transformed average biomass  $\bar{B}$  (panel A), their log-transformed standard  
 648 deviation  $\sigma_B$  (panel B) and their Simpson's concentration index  $\lambda$  (panel C). Black circles, red  
 squares, blue diamonds and green triangles represent respectively 2-, 4-, 8- and 16-species  
 650 mixtures. Solid lines represent the  $y=x$  function.

